

OM of: US-09-805-550-4 to: GenBank: out-format: pfs
Date: Aug 31, 2002 7:35 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

```
-MODEL-firmes.p2n.model -DEV-xlh  
-Q/cgnt2.1/USPTO.spool/US09805550/runal_29082002.160819.375/app-query.fasta.1.897  
-DB-GenBank -OFMT-fastap -SUFFIX-oligop2n.rge -GAPOP-4.500  
-GAPEXT-0.050 -MINMATCH-0.100 -LOOPL-0.000 -LOOPEXT-0.000  
-GAPOP-4.500 -GAPEXT-0.050 -XGAPOP-60.000 -XGAPEXT-60.000  
-FGAPOP-6.000 -FGAPEXT-7.000 -YGAPOP-60.000 -YGAPEXT-60.000  
-DELOP-6.000 -DELEXT-7.000 -START-1 -MATRIX-oligo  
-TRANS-human40.cdi -LIST-45 -DOCALLIGN-200 -THR-SCORE-quality  
-THR_MIN-1 -ALIGN-15 -MODE-LOCAL -OUTFMT-pfs -NORM-ext  
-HEMPSIZE-500 -MINLEN-0 -MAXLEN-200000000  
-USPR-US09805550.ecgnt1.4959 -NCPU-6 -ICPU-3 -LONGLOC  
-DEV-TIMEOUT-120 -WARN-TIMEOUT-30 -NO_XLPHY -WAIT -THREADS-1
```

Search information block:

Query: US-09-805-550-4
Query length: 368
Database: GenBank: *
Database sequences: 1797656
Database length: 187333701
Search time (sec): 3229.260000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP-XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT-XGAPEXT=60.000

Sequence	Strid	Orig	ZScore	EScore	len	Documentation
gb_pat:AR153820	+	368.00	6342.60	0.0	1702	AR153820 Sequence 3 from patent
gb_pat:AX314196	+	56.00	954.35	8.5e-45	353	AX314196 Sequence 7181 from Patent
gb_scs:G71529	-	54.00	920.56	6.5e-43	312	G71529 A94818134FMO17 maize leaf
gb_hlg:AP004214	+	23.00	348.44	4.8e-11	70957	AP004214 Oryza sativa chromosome
gb_pl:AY063103	+	21.00	341.09	1.2e-10	1129	AY063103 Arabidopsis thaliana
gb_pl:AY034912	+	21.00	339.14	1.6e-10	1518	AY034912 Arabidopsis thaliana
gb_pl:AC010793	-	21.00	312.26	5.0e-09	90149	AC010793 Genomic sequence for
gb_hlg:AP003881	-	18.00	258.99	4.6e-06	110775	AP003881 Oryza sativa chromosome
gb_pl:AY037181	-	15.00	235.96	8.8e-05	1378	AY037181 Arabidopsis thaliana
gb_pl:AY058196	+	15.00	235.73	9.1e-05	1426	AY058196 Arabidopsis thaliana
gb_pl:DCRA2311	+	15.00	235.53	9.3e-05	1466	Y12014 Daucus carota mRNA for
gb_pl:FC09	+	15.00	235.41	9.5e-05	1496	Y12014 Daucus carota mRNA for
gb_pl:LES243875	+	14.00	218.10	0.0009	1498	AJ243875 Lycopersicon esculentum
gb_pl:AT4018	+	14.00	191.88	0.0252	80442	AT4018 Arabidopsis thaliana
gb_pat:AR153819	+	13.00	200.69	0.0081	1522	AR153819 Sequence 1 from patent
gb_pl:OS063530	+	13.00	200.65	0.0082	1533	OS063530 Oryza sativa cDNA for
gb_pl:AB005248	+	12.00	157.02	2.20	83646	AB005248 Arabidopsis thaliana
gb_pl:CMS01ATV	+	11.00	171.01	0.3664	720	AT113587 Botrytis cinerea strain
gb_pl:CMS01D80	+	11.00	171.01	0.3664	720	AT113587 Botrytis cinerea strain
gb_in:AR132147	+	11.00	166.28	0.6727	1479	AF132147 Drosophila melanogaster
gb_pl:AY039562	+	11.00	165.43	0.7442	1667	AY039562 Arabidopsis thaliana
gb_in:AF136606	+	11.00	162.32	1.12	2697	AF136606 Drosophila melanogaster
gb_hlg:AC018037	+	11.00	153.02	3.68	11090	AC018037 Drosophila melanogaster
gb_pl:ATAC021640	-	11.00	138.29	24.36	103904	ATAC021640 Arabidopsis thaliana
gb_in:AC010667	-	11.00	135.47	34.99	159570	AC010667 Drosophila melanogaster
gb_in:AC010844	-	11.00	130.88	63.05	320565	AC010844 Drosophila melanogaster
gb_in:LES270958	+	10.00	157.40	2.10	411	AJ270958 Lycopersicon esculentum
gb_in:AF250298	+	9.00	146.11	8.94	165	AF250298 Chironomus balatonicus
gb_in:AF250297	+	9.00	145.91	9.17	170	AF250297 Chironomus plumosus
gb_in:AF250301	+	9.00	145.87	9.21	171	AF250301 Chironomus plumosus
gb_pl:HSZ95373	-	9.00	143.11	10.16	192	HSZ95373 H.sapiens genomic DNA
gb_ba:BSF282971	-	9.00	141.25	16.66	345	BSF282971 Bacterial sp. partial
gb_ba:AF360632	-	9.00	140.31	18.80	398	AF360632 Desulfohalobium sp. C
gb_ba:USU08395	+	9.00	139.04	22.14	483	USU08395 unknown sulfate-reduc
gb_ba:AF121085	+	9.00	138.82	22.76	499	AF121085 uncultured bacterium
gb_ba:USU08394	+	9.00	138.67	23.22	511	USU08394 unknown sulfate-reduc
gb_ba:USU08394	+	9.00	138.67	23.22	511	USU08394 unknown sulfate-reduc

```

|||||
506 ATGGGGCAATCTCTGTGGAGGAATGTGACACATAATTAAC 555
151 GlnleuMetGluMetGlyGlySerThrAspLysAspLysValGln 167
556 CAGCTAATGAGATGGTGGGGCAGTTGGACAAAGATTAAGTCCAAG 605
167 GAlaLeuArgAlaAlaTyrAsnAsnProGluArgAlaValGlyTyrLeu 184
606 GGCTCTCCGTCGGCTTACAAACCCGGAAGCTGCTGTGAATACCTCT 655
184 YrSerGlyIleProValThrAlaGluIleAlaValProIleGlyGln 200
656 ACTCGGATTCAGTACAGCTGAATGCTGTTCCAATGGTGGTCAA 705
201 GAlaAsnThrThrAspArgAlaProThrGlyAlaGlyLeuSerGly 217
706 GGGGAAACACACATGATCGAGCTCTACTGGGAGAGCTGGTCTCTGG 755
217 YIleProAsnThrAlaProLeuAspLeuPheProGlnGlyAlaSerAsn 234
756 GATTCCAACACCGCTCCACTAGATCTTTCGCCGAGGGGCTTCCAATG 805
234 IaGlyGlyAlaGlyGlyGlyProLeuAspPheLeuArgAsnPro 250
806 CTGGAGGTGGTGGTGGTGGACCACTGTGATTTCTTAAGAAACATCA 855
251 GlnPheGlnAlaValArgGluMetValHisThrAsnProGlnIleLeuGly 267
856 CAGTTTCAGACAGTTAGGGAATGTCATCAATCCAAATTTTTCGA 905
267 nPrometLeuValGluLeuSerLysGlnAsnProGlnIleLeuArgLeu 284
906 GCCATGCTGCTGTGAGTTCAGCAAGCAAGCAATCTCAATTTCTAAGTGA 955
284 legGluAsnHisAspGluPheLeuGlnLeuLeuAsnGluProPheGlu 300
956 TTGAGAGCAATCATGATGATGTTCTTCAGTTACTAAATGAGCCCTTTGAA 1005
301 GlyGlyGluGlyAspPheLeuAspGlnProGluGluAspGluMetProH 317
1006 GCGGAGAGGGGGAATTTCTTAGACCAACTGAGAGGATGAATGCCCA 1055
317 SalAlaSerValThrProGluGluGlnGluAlaIleGlyArgLeuGln 334
1056 TGCCATTAGTTTACACAGAGAGAGAGAGCCATTGGAGCGCTTGAGT 1105
334 eRmetGlyPheAspArgAlaArgValIleGluAlaPheLeuAlaCysAsp 350
1106 CCATGGGCTTGACAGACAGCAGCGCTTATTGAAGCAATTTAGCTGCGAT 1155
351 ArgAsnGluGluLeuAlaAlaAsnTyrLeuLeuGlnHisAlaGlyGluGly 367
1156 AGGAACGAGAGAGCTAGACGAACACTATCTCTTGAGCATGCTGGTGAAGA 1205
367 uAsp 368
1206 AGAT 1209

```

seq_name: gb_pat:AX314196

seq_documentation_block:

LOCUS AX314196 353 bp DNA linear PAT 14-DEC-2001

DEFINITION Sequence 7181 from Patent WO0190366.

ACCESSION AX314196

VERSION AX314196.1 GI:17898333

KEYWORDS human.

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (sites) Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

TITLE Human polynucleotides and polypeptides encoded thereby
JOURNAL Patent: WO 0190366-A 7181 29-NOV-2001;
Curagen Corporation (US)
FEATURES location/Qualifiers
source 1..353
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 92 a 71 c 100 g 90 t
ORIGIN

alignment_scores:
Quality: 56.00 Length: 56
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-805-550-4 x AX314196 ..

Align seg 1/1 to: AX314196 from: 1 to: 353

```

253 GlnAlaValArgGluMetValHisThrAsnProGlnIleLeuGlnPro 269
|||||
85 CAGCAGCTTAGGGAATGTCATCAATCCAAATTTTGCAGCTAT 134
269 tLeuValGluLeuSerLysGlnAsnProGlnIleLeuArgLeuGln 286
135 GCTCGTGTGAGTTCAGCAAGCAAGCAATCTCAATTTCTAAGTGTGAGG 184
286 lAsnHisAspGluPheLeuGlnLeuLeuAsnGluProPheGluGly 302
185 AGAATCATGATGATGTTCTTCAGTTACTAAATGAGCCCTTGAAGCGGA 234
303 GluGlyAspPheLeuAsp 308
235 GAGGGGAGTTTCTTTAGAC 252

```

seq_name: gb_sts:G71929

seq_documentation_block:

LOCUS G71929 312 bp DNA linear STS 08-JUN-2001

DEFINITION A94818134FM017 maize leaf DNA Zea mays STS genomic, sequence tagged

ACCESSION G71929

VERSION G71929.1 GI:14333614

KEYWORDS STS.

SOURCE Zea mays.

ORGANISM Zea mays.

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

TITLE Yung, Y.J., Guo, L., Ashlock, D.A., Wen, T.J. and Schnable, P.S.

JOURNAL 3' UTR sequences of maize genes

COMMENT Unpublished

Contact: Schnable, P.S.

Schnable laboratory

Iowa State University

GA05 Agronomy Hall, Ames, IA 50011, USA

Tel: 515-294-0975

Fax: 515-294-2299

Email: schnable@iastate.edu

Primer A: GGAGACCAACCGTAGC

Primer B: TTAGACCAACCTGAGAGG

PCR Profile:

Denaturation: 94 degrees C for 30 seconds

Annealing: 60 degrees C for 45 seconds

Polymerization: 72 degrees C for 90 seconds

PCR cycles: 31

Thermal cycler: Perkin Elmer TC

Protocol:

Template: 10-20 ng

Primer: each 0.5 uM
 dNTPs: each 200 uM
 Taq Polymerase: 0.05 units/uL
 Total vol: 20 uL

Buffer:

MgCl₂: 2 mM
 KCl: 50 mM
 Tris-HCl: 20 mM
 pH: 8.4

FEATURES
 source 1. 312
 Location/Qualifiers

/organism="zea mays"
 /strain="DE811"
 /db_xref="taxon:4577"
 /clone_lib="maize leaf DNA"
 /note="PCR products amplified from genomic DNA"

STS
 BASE COUNT 86 a 82 c 54 g 90 t
 ORIGIN

alignment_scores:

Quality: 54.00 Length: 54
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-805-550-4 x G71929/rev ..

Align seg 1/1 to reverse of: G71929 from: 1 to: 312

315 MetProHisAlaIleSerValThrProGluGluGluGluAlaIleGlyAr 331
 |||||||
 311 ATGCGTCATGCATTAGTGTACACAGAGAGAGAGAGCGCATTTGACG 262
 |||||||
 331 gLeuGluSerMetGlyPheAspArgAlaArgValIleGluAlaPheLeu 348
 |||||||
 261 GCTTGAGTCCATGGGTTTCACAGACGCGTTATCGAAGCATTCCTAG 212
 |||||||
 348 lacysAspArgAsnGluGluLeuAlaIleAsnTyrLeuLeuGluHisAla 364
 |||||||
 211 CCTCGCATAGAGACGAGAGAGCTAGACCAACTCTCTCTTGACATGCT 162
 |||||||
 365 GlyGluGluAsp 368
 |||||||
 161 GGTGAGGAGAGAT 150
 |||||||

seq_name: gb_htg:AP004214

seq_documentation_block:

LOCUS AP004214 70957 bp DNA linear HTG 27-SEP-2001
 DEFINITION Oryza sativa chromosome 2 clone OJ1118_B07, *** SEQUENCING IN
 PROGRESS ***, in ordered pieces.

ACCESSION AP004214 GI:15787871
 VERSION AP004214.1 GI:15787871
 KEYWORDS HTG, PHASE2.
 SOURCE Oryza sativa (cultivar: Nipponbare) DNA, clone: OJ1118_B07.
 ORGANISM Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE
 AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
 TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC
 clone: OJ1118_B07

JOURNAL
 REFERENCE Published only in Database (2001) In press
 AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
 TITLE Direct Submission

JOURNAL Submitted (26-SEP-2001) Takuji Sasaki, National Institute of
 Agricultural Sciences, Rice Genome Research Program, Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan

COMMENT

(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
 Tel: 81-298-38-7441, Fax: 81-298-38-7468)
 The nucleotide sequence of this BAC clone was generated by
 combining Monsanto and RGP-Japan sequencing data.
 NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is
 available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES

source

1. 70957
 /organism="Oryza sativa"
 /cultivar="Nipponbare"
 /db_xref="taxon:4530"
 /chromosome="2"
 /clone="OJ1118_B07"

BASE COUNT 20235 a 14617 c 14676 g 21429 t
 ORIGIN

alignment_scores:

Quality: 23.00 Length: 23
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-805-550-4 x AP004214 ..

Align seg 1/1 to: AP004214 from: 1 to: 70957

163 AspIysValGlnArgAlaLeuArgAlaIleArgAsnAsnProGluAlaGly 179
 |||||||
 46582 GATAAAGTCACAAAGACCTCTACGCTTAACCAACCCAGACGCTGC 46631
 |||||||
 179 aValGluTyrLeuTyrSer 185
 |||||||
 46632 TGTTCGAATATCTATATTCT 46650
 |||||||

seq_name: gb_pl:AY063103

seq_documentation_block:

LOCUS AY063103 1129 bp mRNA linear PLN 26-NOV-2001
 DEFINITION Arabidopsis thaliana putative RAD23 protein (F20B17.8) mRNA,
 complete cds.

ACCESSION AY063103
 VERSION AY063103.1 GI:17104776
 KEYWORDS FLI CDNA.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
 AUTHORS Yamada, K., Banb, J., Banno, F., Chang, E., Dale, J.M., Goldsmith, A.D.,
 Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C., Toriumi, M., Wu, H.C.,
 Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H.,
 Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A.,
 Karlin-Neumann, G., Kawai, J., Kim, C., Koeseema, E., Lam, B., Lin, J.,
 Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,
 Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,
 Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Full Length cDNA of gene F20B17.8 (GI:1715605)
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1129)
 AUTHORS Yamada, K., Banb, J., Banno, F., Chang, E., Dale, J.M., Goldsmith, A.D.,
 Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C.,
 Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H.,
 Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A.,
 Karlin-Neumann, G., Kawai, J., Kim, C., Koeseema, E., Lam, B., Lin, J.,
 Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,
 Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,
 Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

REFERENCE

JOURNAL
 REFERENCE 2 (bases 1 to 1129)
 AUTHORS Yamada, K., Banb, J., Banno, F., Chang, E., Dale, J.M., Goldsmith, A.D.,
 Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C.,
 Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H.,
 Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A.,
 Karlin-Neumann, G., Kawai, J., Kim, C., Koeseema, E., Lam, B., Lin, J.,
 Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,
 Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,
 Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

TITLE
JOURNAL

COMMENT

Karlin-Neumann, G., Kawai, J., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
Direct Submission
Submitted (15-NOV-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the pN1 (ORF) clones using the RAFL cDNAs: Yamada, K., Banh, J., Banno, F., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shim, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

FEATURES
source

location/Qualifiers

1..1129

/organism="Arabidopsis thaliana"

/db_xref="taxon:3702"

/chromosome="1"

/clone="U09913"

/note="This clone is in pN1 51. ecotype: Columbia"

1..1129

/gene="F20B17.8"

1..1098

/gene="F20B17.8"

/note="contains two UBA PF00627 domains and is a member of the ubiquitin PF00240 family"

/codon_start=1

/evidence=experimental

/product="putative RAD23 protein"

/protein_id="FALJ4277.1"

/db_xref="GI:17104777"

/translation="MKLVKTLGSHFEIVLPSDTIMAVKNIEDSOGKDNYPGCGQ

LT.IHNGVLNDETSLEKNKTEGFLVNLKSKSGSAQASQVPSATTSRPPAA

PSTVQSSVPASPIPAQEQPAQDTGQAATLVSSSLEMQOQIMEGSGMSDE

TTRALRAAYNNPRAVDLYSGIPQAEVAVPPEAIOIGSGAAPAPASGGGNSSP

LDLPEQETVAAGSGDLGTEFLRNDDFOOLRTMWHNSPOIIOQPMCELEKQNPOL

RLIOENAEFLQVNEPYEGSDGDFDPDQEMPHAINVTPEQEAIORELEMGFD

RALVTEAFILACDRNEELANNTLENSGDFED"

misc_difference

576

/gene="F20B17.8"

/note="compared to genomic sequence"

/replace="g"

1099..1129

/gene="F20B17.8"

BASE COUNT 332 a 254 c 261 g 282 t

ORIGIN

Alignment_scores:

Quality: 21.00 Length: 21
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment_block:

US-09-805-550-4 x AY063103 ..

Align seg 1/1 to: AY063103 from: 1 to: 1129

342 VALLIAGLUAAlaPheLeuAlaCysAspArgAsnGluGluLeuAlaAlas 358
|||||

1012 GTAAATGAAGCGCTTCTCCGATGTGACCGAATGAAGAAATTGGCAGCTAA 1061
358 nryrleuLeuGlu 362
|||||
1062 CTATCTATGTGAG 1074

seq_name: gb_pl:AY034912

seq_documentation_block:

LOCUS AY034912

DEFINITION Arabidopsis thaliana putative RAD23 protein (F20B17.8) mRNA.

complete cds.

ACCESSION AY034912

VERSION AY034912.1 GI:14334441

KEYWORDS F11 cDNA.

SOURCE

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1518)

REFERENCE

AUTHORS

Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K.,

Goldsmith, A.D., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Yu, G.,

Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y.,

Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,

Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M.,

Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M.,

Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and

Theologis, A.

Full length cDNA of gene F20B17.8 (GI:7715605)

2 (bases 1 to 1518)

Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K.,

Goldsmith, A.D., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Yu, G.,

Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y.,

Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,

Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M.,

Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M.,

Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and

Theologis, A.

Direct Submission

Submitted (10-MAY-2001) Plant Gene Expression Center, 800 Buchanan

Street, Albany, CA 94710, USA

RIKEN Genomic Sciences Center (GSC) members carried out the

collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN

Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,

Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,

Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the

sequencing and annotation of the RAFL cDNAs: Yamada, K., Liu, S.X.,

Sakano, H., Pham, P.K., Banh, J., Chung, M.K., Goldsmith, A.D.,

Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Yu, G., Bowser, L.,

Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Koesema, E.,

Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J.,

Shim, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to

this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC)

contributed equally to this work as PIs.

FEATURES
source

location/Qualifiers

1..1518

/organism="Arabidopsis thaliana"

/db_xref="taxon:3702"

/chromosome="1"

/clone="RAFL05-08-C09 (R09913)"

/note="This clone is in a modified pBluescript vector (lambda zap) as a XhoI/SstI insert.

ecotype: Columbia"

1..136

/gene="F20B17.8"

1..1518

/gene="F20B17.8"

5' UTR

gene

```

CDS
    137..1234
        /gene="F20B17.8"
        /note="contains two UBA PF00627 domains and is a member of
        the ubiquitin PF00240 family"
        /codon_start=1
        /evidence=experimental
        /product="putative RBD23 protein"
        /protein_id="AAK59419.1"
        /db_xref="GI:14334442"
        /translation="MKLVTKLKGSHFEIRVLPSTIMAVKNKIEDSGKNDYPCGQ
        LLINQKVLKDEPTSEIVKRVKEEGLVWLSKSGSGAGASVQPVATTSSTPAA
        PSTOSSPVASPIPAOQPAOATDTYGOAASLYSGSSLEOMQIEMGSGSKDE
        TVTRALRAVNNPERANDYLSGIPQTRAVAVPPPAIAGSAGAPVAPAGGSPSP
        LDLPQETVAAGSGDGLTEFLRNNDQEQQLRTVWHSNPQILOPMQELGKQNPQL
        RLIDNQAEPLOLVNEPEYSGDGEKMDPEDEQEMPHAINVTPEQEAIORLEAMGFD
        RALVLEAFVACDRNEELIANYLENSGDFD"
        712
    misc_difference
        /gene="F20B17.8"
        /note="compared to genomic sequence"
        /replace="g"
        1235..1518
        /gene="F20B17.8"
    3'UTR
        /gene="F20B17.8"
    misc_difference
        /gene="F20B17.8"
        /note="not present in genomic sequence"
    BASE COUNT      421 a      326 c      338 g      433 t
    ORIGIN

alignment_scores:
    Quality: 21.00      Length: 21
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
    US-09-805-550-4 x AY034912 ..

Align seg 1/1 to: AY034912 from: 1 to: 1518

    342 ValIIGLlUAlaPheLueuAlaCyAsPaRGaNGlUGlUeLueAlaLaas 358
    |||||||
    1148 GTATATGAACCGTCTCGCATGTGACGAAATGMAAGAAATTGGCAGCTAA 1197

    358 nTYrLeuLeuGlu 362
    |||||||
    1198 CTATCTATTGGAG 1210

seq_name: gb_pl:AC010793

seq_documentation_block:
    LOCUS      AC010793      90149 bp      DNA      linear      PLN 07-MAY-2000
    DEFINITION  Genomic sequence for Arabidopsis thaliana BAC F20B17 from
    chromosome I, complete sequence.
    ACCESSION  AC010793
    VERSION    AC010793.3      GI:6938146
    KEYWORDS   HTG.
    SOURCE     thale cress.
    ORGANISM   Arabidopsis thaliana
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
    1 (bases 1 to 90149)
REFERENCE
    AUTHORS    Khan,S., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Kim,C.,
    Shinn,P., Altafi,H., Bel,O., Chin,C., Chlou,J., Choi,E., Conn,L.,
    Conway,A., Gonzales,A., Hansen,N., Howng,B., Koo,T., Lam,B.,
    Lee,J., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharsky,N.,
    Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,T., Soultwick,A.,
    Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Federspiel,N.A.,
    Theologis,A. and Ecker,J.R.
    TITLE      Genomic sequence for Arabidopsis thaliana BAC F20B17 from
    chromosome I
    JOURNAL    Unpublished
    REFERENCE  2 (bases 1 to 90149)

```

```

AUTHORS      Ecker,J.R.
TITLE        Direct Submission
JOURNAL      Submitted (23-SEP-1999) Arabidopsis thaliana Genome Center,
              Department of Biology, University of Pennsylvania, 38th Street and
              Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
              3 (bases 1 to 90149)
REFERENCE
    AUTHORS    Ecker,J.R.
    TITLE      Direct Submission
    JOURNAL    Submitted (08-FEB-2000) Arabidopsis thaliana Genome Center,
              Department of Biology, University of Pennsylvania, 38th Street and
              Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
              4 (bases 1 to 90149)
REFERENCE
    AUTHORS    Cheuk,R., Shinn,P., Brooks,S., Buehler,E., Chao,Q.,
    Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bel,B., Chin,C.,
    Chlou,J., Choi,E., Conn,L., Conway,A., Gonzales,A., Hansen,N.,
    Howng,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J.,
    Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H.,
    Schwartz,T., Soultwick,A., Thaveri,A., Toriumi,M., Vaysberg,M.,
    Yu,G., Davila,R., Federspiel,N., Theologis,A. and Ecker,J.
    TITLE      Direct Submission
    JOURNAL    Submitted (07-MAY-2000) Arabidopsis thaliana Genome Center,
              Department of Biology, University of Pennsylvania, 38th and
              Hamilton Walk, Philadelphia, PA 19104-6018 USA
    COMMENT    On Feb 8, 2000 this sequence version replaced gi:6453825.
    FEATURES
    source
    Location/Qualifiers
        1..90149
            /organism="Arabidopsis thaliana"
            /db_xref="taxon:3702"
            /chromosome="1"
            /clone="F20B17"
            complement(join(3669..4322,4410..4681,5385..5574))
            /note="hypothetical protein"
            /codon_start=1
            /evidence=not_experimental
            /product="F20B17.1"
            /protein_id="AAF68129.1"
            /db_xref="GI:7715611"
            /translation="MEIGSSSTVAGGGLSVPPGRRFRTEELLYLKKRYSEPI
            DLDVREVDLNLKEPEWELKEKCRIGSGPQNEVYFSSDKKYPGTPTNRRAATAGFWK
            AGRDCKSHLNSKKIGLRKTLVFTYTGAPRHOKEKHEWIMHEYRLDSSNEISLOEOWVY
            CRVFKKHNHFRGEHOERODHNNHNYISTNDDHNNHNDHNSNSRPLILRPHDH
            HHHNHGRIHRLPRANTLSHSGMHPQLPESPSSAAAAAASAPFPYSPITTD
            IECSONLRLTISNNYGDWSEFLDKLTLTGNNQDQDQVQNHQAKCGDLSNNDND
            QADHIGNNNGSSSPVNOREFHYHLDANLKRPK"
            join(9713..9930,10050..10255,10351..10577,10822..10953)
            /note="unknown; similar to ESTs gb|T04661.1, and
            emb|235002.1"
            /codon_start=1
            /evidence=not_experimental
            /product="F20B17.2"
            /protein_id="AAF68106.1"
            /db_xref="GI:7715588"
            /translation="MASSSDPMWRREYNRAIKSIDINGMMSRNASGILTPDAORRAS
            ATRKTTILGTRLDLSLVLKVPGRKQVSEKMRKRDVGNLRSKTQVASALNMS
            NFANDSLFTGDLKRPDAINRVSGMDNGIVFOQVAKEDQEGLEKLEYVMSYKHI
            ALAVNEELTQTRLIYVLCRCISLFLILISVSNMGSSMODDIDYVDITDSRLRR
            VOKSLAMKNSKSSGSCMSMLSVGLVGLVIMLVKYL"
            complement(join(12754..12882,12967..14412,14530..14757,
            14833..15165))
            /note="unknown; similar to ESTs emb|230470.1,
            gb|A193698.1, gb|N96085.1, and emb|230471.1"
            /codon_start=1
            /evidence=not_experimental
            /product="F20B17.3"
            /protein_id="AAF68128.1"
            /db_xref="GI:7715610"
            /translation="MSLVVGGSLGLTVGDLSLNSTKNWKSFFSVNRRRLRAA
            LVQAPKEDGAASPSPSPASVVOYRRADLADLOAEARLGAIDASISYSPDLA
            RKHGQSPKALRSLRSLIGALGFAKLIQIOKQNLKNNMKRAIEIRLFTPRGPT
            FVKLGQSTPRDLCPPLYLELAELOALPFPAPAEACIERLDSLETFPSVS
            PPPIAASLGOVYKAOIRYSGOVAVVAKVORCIEARLIGDYLKNGVCKLTKKYVDFI
            TTDTVTLIDFRACRYOELNVOEONARRKKLYADKADYLVPIFDVYSRVKVTM
            EWEGTKLNEOLAIESGILKYLIVNTIGTQSLRQLLEYGFFHADPHGNNLATPDGK

```

CDS

LAFLDFGMSSETEPEAFIIGHVHLVNDYEMANDYALAKFLSPDVTPIPL
RDFPDDALNTVSELNFKTLVDGLGAVFYQFENVYVALILKSLVLESLAYAP
NFKVLAASYPAKRLITDNPRLIDELPFGFRNNRLNLELQOOSKXDPA
KDALQPVKLTLDPNGEELRLVKEAVRSEALATVVDYTLSEPLRSLVFN
GNGPLSTAELOSTELRDVSRWGLQSSSEFPAIIQPIQVLAQOPEARLGR
VAGVGRLARARLQILRRATPSSASP

complement(join(15862, .15982, 16100, .16200, 16324, .16407,
16341, .16832, 17100, .17144, 17245, .17342, 17518, .17678,
18130, .18224, 18308, .18384, 18435, .18497, 18587, .18698,
19119, .19194, 19351, .19434, 19533, .19719))
/note="similar to sodium-hydrogen exchanger NHE1
(AF142676)"

CDS

/codon_start=1
/evidence=not_experimental
/product="F20B17.4"
/protein_id="AAE68127.1"
/db_xref="GI:7715609"
/translation="MSSELOISPAIHDPGQEQOAGVILLOIMVLVSLGVH
LRHKPYLLEPASLLIGLVGLANLSTERTSIRVEFLISFRHGSISTSSSE
CFCLPSTYILKLEYLGVMFLRLFEVCELMGLSATDPTVLSITQELSDVN
LYALVESEVLNDADIEVTLIRSFSLCQWMAISLYRTMSLVRSMSGONFVYI
VRLFEVGSMSAMKFIYLLSILSVRTWSAVSSYFPHSHKTLLEYTSVSIY
FTLIEIYQFMKHTYNSLSANSOREFAEFHLISSIAETFEVFIYMGEDLAEMHSA
ANYEGGCIYVNLARPAHRTIPMTHOKALMYSGKTLICVPLSYGCVSVNTKCGCG
IGLRGMAAFALAQSVHDEPEGHQITFTATATVTVLLIGSTGTMLEALYVGD
SHDLSLGDGEVNSRTYSTDDETPPGSGFRKLEFHKSAASFELDRNYLTPEF
TSNNGDDDEGMEQHGNHNL"

complement(join(20282, .20788, 20865, .21100, 21175, .21394,
21491, .21630, 21764, .21937, 22013, .22701, 22782, .22853,
22939, .23013, 23095, .23163, 23251, .23322, 23560, .23634,
23717, .23878, 23973, .24044, 24146, .24220, 24312, .24383,
24506, .24629, 24713, .24823))
/note="unknown; similar to Ests gb|A197133.1"
/codon_start=1
/evidence=not_experimental
/product="F20B17.5"
/protein_id="AAE68126.1"
/db_xref="GI:7715608"
/translation="WGSNTVATSRLLICFASFTVFSNMISSVTPDRDAALRSIM
DNDNTPPSGSDPCGTPEGVSCNNSHITAGLSTGKRLSGDIELAELRL
DLSPNRLTGLSLRGLDLOKRLNLLIAGCGFTGTIPNEGLYKDLSPFLANSNFG
KIPASGLNTRKWLADLADNOLGPIPISSGSSGDLILKAKHFRNKQLOSTIPP
KLFSSEMILIHVRLDRMTLGVKPEPENSNTITIELANLKNKYSKPLSDMSMAY
VLSNNSFDESEPLRSTPLSLTYMEGSLGPIPKLFGPQLOVRLKNAAN
GLSLGDTGVEQLVDLQNDLSSVLSGYNLWYLSFELNVSOLYODIKRA
LSCCRLEGNVCTTALSNYCOIQOQVAKRYSTSLANGCKSPDOQVPSOCC
AYVEGTLYFRGPMFRLSNVNTYHSLMSLWGLTLPGSVSLQNPFNNDYLOIO
LALFPPMGKXENREYORIGEDLSNOTKPPLEGPYFTASPTFPADNGHLSR
MWTGTCGSAIVCALGITYAMMOKRARAAGTSGSKSVTLVSSASSKDSG
GAPOUKARWFSTYELKTKTNNSVSSELTGIGKTKMLDGHVATKRAOQST
OGLEFTELELSRVHKNLVGLVGCFFQEGQILVLEYMSNGLSDSLTRSGITL
DMKRLVALAGSARGLAYLHLDAPPIIHDDVSTNLIDENLTAQVADGLSKLVSD
CTKGHVSQVKGTLGYLDPEYTTQKLTKESDVYSFVVMMLTAKOPLKGYIVR
EIKLVNKSDDDFGLRDKDRSLRDVGTLPGLGRYELALKCVDTEADREPTSEVY
KEIETIIONGASSSSASASSATDGEKLLVGGTAKKKEARDGDCGAFDYSGGIS
VPKIEPK"

CDS

complement(join(25580, .26059, 26154, .26582, 26669, .27097,
27555, .27818))
/note="hypothetical protein"
/codon_start=1
/evidence=not_experimental
/product="F20B17.6"
/protein_id="AAE68125.1"
/db_xref="GI:7715607"
/translation="MGLCYSVDRITGKEPEASATATTAETVEERSSGGRMRPDK
GGDIBGIPOLVGLVNSGSKTACTOOGKGTNDADLVENFESRDPDYCGVF
DGHGPRGHMAYAKVRDILPPTLLTQKMTSESOSSTVANGROIKFCEEEVOTTS
EQVKTESVTIMDEQCELPNVNNDLPMITLPLKAMKASCOQIDKEIKMHTTDC
FCGGTISVTLIKOGEDLVGNIGDSRAVLATREDNALAVQITLIDKPLGLTLOI
MLNSERSQKQELNFRFLCVLSGESARVLCOKGRVALLDEPVAVMLPNSPGL
AMARAFDCLKGLYSVDIPVIRLTEREDOFIILASDVMLDYSKVEKVDIYASAP
SRSTAAALVDYAVRSWIRIKYPTSKNDKCVCLTODSSVAMVSVSKKVDSEES
IESVTNSTSKEDDEIYVRKDEKIPESGIESKMTATLACISVADDERMSALEGIT

CDS

RVNSLSIPRFLSGELNRSWKWL"
complement(join(29562, .29631, 29861, .29925, 30172, .30291,
30565, .30659, 30757, .30869, 31083, .31133, 31431, .31498,
31643, .31706, 31868, .31893, 32190, .32360, 32589, .32670,
32937, .33061, 33213, .33467, 33581, .33739))
/note="hypothetical protein"
/codon_start=1
/evidence=not_experimental
/product="F20B17.7"
/protein_id="AAE68124.1"
/db_xref="GI:7715606"
/translation="MERKKYPIDGHEHYTYEFGQGSALVHRLCPDENVAKTL
DPERNDCLDNISREAOITMLVDPNLIKSHCSFVSHNMLVPIHSGSCCHITKA
AYPDGEFEALITATLREALGLDYLHGHGHIHRVAKGNLILARAGVAKLGDGVSC
LFDSGDRQRTNRTFVGTCPWADWISFQITGIELAHGAPAPSKYPPKPSFOMIASCLY
KDSKRSKAKKLHSPFKOARSDDYIARLKLDELPLVNRVOAKINEYKRGISGLSG
SATSIALODSDOSEIOEDPGITNNYLOPLHRSLSIARDSDDSDSLASPSYDY
VYSPRHEDSLNTHVSGTHANNGKTDKSTPTNPTETILINGSVLAQGNAPNG
ESDTQOQLONGSCKNTHFPVGGDDVPTEMAKVPPKASSLSDSDSKRPVQOQNG
RFKVTSENLDIEKVLCOHSSASLPHSDVTLNLTSSVYLPVLPVLTNLERDVOL
NKEVYNNILCPOLNPMWLICAEELQKTEHQAQVSMN"

CDS

complement(join(35055, .35171, 35264, .35348, 35471, .35580,
35968, .36075, 36158, .36283, 36360, .36530, 36765, .36828,
36929, .36984, 37041, .37079, 37300, .37461, 37716, .37811))
/note="similar to Rad23, isoform I (F12013); similar to
Ests gb|T42959.1"
/codon_start=1
/evidence=not_experimental
/product="F20B17.8"
/protein_id="AAE68123.1"
/db_xref="GI:7715605"

alignment_scores:

Quality:	21.00	Length:	21
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-805-550-4 x AC010793/rev ..

Align seg 1/1 to reverse of: AC010793 from: 1 to: 90149

342 valllieglualaphelenua1acyasapagaangluiglu1enua1a1aas 358
|||||
35141 GAAATAGGAAGCCCTTCCTCCATGTGACCGAAATGGAATTCGACGCTAA 35092

358 ntYrleu1enua1u 362
|||||

35091 CTATCTATTGGAG 35079

seg_name: gb_htg:AP003881

seg_documentation_block: 110775 bp DNA linear HTG 10-JUL-2001

LOCUS AP003881 110775 bp DNA linear HTG 10-JUL-2001
DEFINITION Oryza sativa chromosome 8 clone OJ1124_B05,*** SEQUENCING IN
PROGRESS ***, In ordered pieces.

ACCESSION AP003881
VERSION AP003881.1 GI:14646814
KEYWORDS HTG; HTGS PHASE2.
SOURCE Oryza sativa (cultivar: Nipponbare) DNA, clone: OJ1124_B05.
ORGANISM Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ethnartoidae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 110775)
Sasaki, T., Matsumoto, T. and Yamamoto, K.

AUTHORS Oryza sativa nippobare(GA3) genomic DNA, chromosome 8, BAC
TITLE clone: OJ1124_B05
JOURNAL Published Only in Database (2001) in press

REFERENCE 2 (bases 1 to 110775)
Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2001) Takuji Sasaki, National Institute of

Agrobiological Resources, Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsasakib@r.affrc.go.jp, URL: http://rpg.dna.affrc.go.jp/, Tel: 81-298-38-7441, Fax: 81-298-38-7468)
The nucleotide sequence of this BAC clone was generated by combining Monsanto and RGP-Japan sequencing data.

NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES

source
1..110775
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="8"
/clone="OJ1124.B05"
BASE COUNT 31727 a 23364 c 23821 g 31760 t 103 others
ORIGIN

alignment_scores:

Quality: 18.00 Length: 18
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-805-550-4 x AP003881/rev ..

Align seg 1/1 to reverse of: AP003881 from: 1 to: 110775

167 ArgAlaLeuArgAlaAlaTyrAsnAsnProGluArgAlaValGluTyrLe 183
|||||
86919 CGAGCTCTGGAGCTGCATACACATCAGAACGCGCTGTGGAGTATT 86870

183 UTYR 184
|||||

86869 GTAC 86866

seq_name: gb_pl:AY037181

seq_documentation_block:

LOCUS AY037181 1378 bp mRNA linear PLN 10-JUN-2001
DEFINITION Arabidopsis thaliana unknown protein mRNA, complete cds.
ACCESSION AY037181
VERSION AY037181.1 GI:14335003

KEYWORDS
FLI CDNA.

SOURCE

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1 (bases 1 to 1378)

REFERENCE

AUTHORS

TITLE

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1378)

AUTHORS

Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Tracy, S.E., Ban, H., Bowser, L., Carninci, P., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Tortum, M., Yamada, K., Yu, G., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
Arabidopsis cDNA clones

Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Tortum, M., Yamada, K., Yu, G., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

TITLE

JOURNAL

COMMENT

Direct Submission
Submitted (22-MAY-2001) Salk Institute Genomic Analysis Laboratory (Signal), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RFLP cDNAs (RFLP cDNA: "RIKEN Arabidopsis Full-Length cDNA") : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PEGC (SSP) Consortium members carried out the sequencing and annotation of the RFLP cDNAs: Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Tracy, S.E., Ban, H., Bowser, L., Chung, M.K., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Palm, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M., Pham, P.K., Pham, P.K., Quach, H.L., Sakano, H., Southwick, A., Tang, C.C., Tortum, M., Yamada, K., Yu, G., Davis, R.W., Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

source

1..1378
Location/Qualifiers
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="5"
/clone="RAFL09-12-L08(R12657)"
/note="ecotype: Columbia"
1..70
71..1207
/codon_start=1
/product="unknown protein"
/protein_id="AAK59766.1"
/db_xref="GI:14335004"
/translation="MKIFVKTLSGSNFELEVKPADKVSQVKALETWKAEPAAKQMLHOGKVLKDETTLENNVENSFTVIMSKTASPSGSTASAPASATOPQVATPQVSAPTSVPTSGTATAAADATASVOTDYGQAAMVAGTLESTVOOILDMGCGSMRDRTVIRALRAAFENPERAVEYLSGIPQAEIPLPAQAPATGEQANLADPQAAAPAAATGGPNANPLNPFPCPMADAGAGNIDFLNESHOPALRAMVQANQIIOPMIOELGKQNPQVRLIOEHADPFLRLNVEYEGEENMEQLEAAMPQAVITFEERDAIERLEGGFDRAMVLEVFACNKNKEELANLLDHMEFEDQ"

5'UTR
CDS

3'UTR
BASE COUNT 366 a 332 c 296 g 384 t
ORIGIN

alignment_scores:

Quality: 15.00 Length: 15
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-805-550-4 x AY037181 ..

Align seg 1/1 to: AY037181 from: 1 to: 1378

174 AsnAsnProGluArgAlaValGluTyrLeuTyrSerGlyIlePro 188
|||||
599 AACATCTCTGAAGAGCTGTGAATACCTGTACTCGGATATCCCT 643

seq_name: gb_pl:AY058196

seq_documentation_block:

LOCUS AY058196 1426 bp mRNA linear PLN 04-NOV-2001
DEFINITION Arabidopsis thaliana unknown protein mRNA, complete cds.

ACCESSION AY058196

VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AY058196.1 GI:16648837
FLI_CDNA.
thale cress.
Arabidopsis thaliana
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1426)
Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
Arabidopsis cDNA clones
Unpublished
2 (bases 1 to 1426)
Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
Direct Submission
Submitted (28-AUG-2001) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN
Arabidopsis Full-Length CDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

COMMENT

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H.,
Kim,C.J., Koesema,E., Meyers,M.C., Shin,P., Banh,J., Bowser,L.,
Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G.,
Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A.,
Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,
Davis,R.W., Theologis,A. and Ecker,J.R.

FEATURES

Source

1. 1426
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="5"
/clone="RAFL09-10-108(R14049)"
/note="ecotype: Columbia"
1. 78
79. 1215
/codon_start=1
/product="unknown protein"
/protein_id="AA125609.1"
/db_xref="GI:16648838"
/translation="MKLFVTKLSGSNEIEVFPADKYSVDTATLETYKGAEPRAKQK
LIHQGVLKDETLLENVNVSEFIVMLKTRKPSGASTASAPASATOPOTVAPR
QVAPFASVPTSGATAPATPAASVOTVGOAASNLVAGTTLESTVOYLDMGK
GSMDDTVVVALRAAFNNPRAVEYLSGIPDAEIPVPAAGATPGGOAANPLAOPD
AAPAATGCPNANPLNLFPGQHPADAGACAGUDPLRSOQALIRANVOANPOL
OPMDELGKQNPOLVRLIOEHQDFRLINIEPVEGEEVNEQLEAANPAQATVTPER
EAIERLEGKGFDRAMVLEVFACKNKEELNAYLLDHHEFDD"

5' UTR
CDS
3' UTR
BASE COUNT
ORIGIN

377 a 343 c 301 g 405 t
1216 .1426

alignment_scores:
Ratio: 15.00 Length: 15
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-805-550-4 x AY058196 ..
Align seg 1/1 to: AY058196 from: 1 to: 1426

174 AASAPProGUARGAlaValGluYrLeuYrSerGlyIlePro 188
|||||
607 AACAAATCTGAAGAGCTGTGAATACCTGATCGGATCCCT 651

seq_name: gb_pl:DCRAD2311

seq_documentation_block:
LOCUS DCRAD2311 1466 bp mRNA linear PLN 28-JAN-1999
DEFINITION Daucus carota mRNA for RAD23 protein, isoform II.
ACCESSION Y12014
VERSION Y12014.1 GI:1914684
KEYWORDS isoform II; RAD23 (ubiquitin-like protein involved in UV excision repair).
SOURCE carrot.
ORGANISM Daucus carota
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; eusterids II; Apiales; Apiaceae; Daucus.
REFERENCE 1 (bases 1 to 1466)
AUTHORS Sturm,A. and Lienhard,S.
TITLE Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant
JOURNAL Plant J. 13 (6), 815-821 (1998)
MEDLINE 98345997
REFERENCE 2 (bases 1 to 1466)
AUTHORS Sturm,A.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-1997) A. Sturm, Friedrich Miescher-Institut,
Maulbeerstr. 66, CH-4058 Basel, SWITZERLAND

FEATURES

Source

1. 1466
/organism="Daucus carota"
/isolate="W01C"
/sub-species="Queen Anne's lace"
/db_xref="taxon:4039"
40. 1179
/function="assembly factor of the complex for nucleotide
excision repair of V-damaged DNA"
/codon_start=1
/product="RAD23 protein, isoform II"
/protein_id="CAA72742.1"
/db_xref="GI:1914685"
/db_xref="SPRMBL:003991"
/translation="MKLFVTKLSGSHPIRAQPNQDYMAIKKNIDILQGDKNPCCGO
LLIHNGVLDKDESTLAKSKISEDFLVVLMKSKTMSSTGTPAAGSSAPAPAPAY
APAPAPAAPASAVIPNTTVPEAPLSFAFASDYVGEASNVVAGSNLEGTIHKM
GGGMQNTNVSRAIRAYNNPERAVDLYSGIPMAEAVVSHFGQDQINAGNNAT
DNGVGAAPAGPNSLPIPLMFPEOFTLSGVTGAGCSLEFLNNPQFQRLRBMVRNPQI
LQPMLELQKQNPOLLQIOEHHEFRLQILNERPEASBEGCMFDQPEDDVQELITVTA
DQEAIERLEAMGDFRGVLTIEAFKACDNEELNAYLLNAGDFDD"

CDS

BASE COUNT 405 a 311 c 326 g 424 t

ORIGIN

alignment_scores:
Ratio: 15.00 Length: 15
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-805-550-4 x DCRAD2311 ..

Align seg 1/1 to: DCRAD231I from: 1 to: 1466

342 ValIleGluAlaPheLeuAlaCysAspArgAsnGluGluLeuAla 356
 1093 GTTATTGAAGCCCTTTTACCATGTGACCCGACGAGGAATTGGCG 1137

seq_name: gb_pl:DCRAD231

seq_documentation_block:

LOCUS DCRAD231 1496 bp mRNA linear PLN 28-JAN-1999

DEFINITION Daucus carota mRNA for RAD23 protein, isoform I.

ACCESSION Y12013

VERSION Y12013.1 GI:1914682

KEYWORDS isoform I; RAD23 (ubiquitin-like protein involved in UV excision repair).

SOURCE carrot.

ORGANISM Daucus carota

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; easterids II; Apiales; Apiaceae; Daucus.

REFERENCE 1 (bases 1 to 1496)

TITLE Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in yeast

JOURNAL Plant J. 13 (6), 815-821 (1998)

MEDLINE 98345397

REFERENCE 2 (bases 1 to 1496)

AUTHORS Sturm, A.

TITLE Direct Submission

JOURNAL Submitted (24-MAR-1997) A. Sturm, Friedrich Miescher-Institut, Maulbeerstr. 66, CH-4058 Basel, SWITZERLAND

FEATURES

source

1..1496

/organism="Daucus carota"

/isolate="W001C"

/sub-species="Queen Anne's Lace"

/db_xref="taxon:4039"

63..1211

/function="assembly factor of the complex for nucleotide excision repair of V-damaged DNA"

/codon_start=1

/product="RAD23, isoform I"

/protein_id="CAI72741.1"

/db_xref="GI:1914683"

/translation="MKIYVKILKSGSEIOYVNPDSADVVKRSTETAGAAVYPAAO

MLIYQKYLKQKTLLENNAVENSTIVIMLSKSPSEGSTTSTRAAPKAPQTSAPP

SVPAVAVSQPASTLPVPAPSPAPAPATAPIPSAVGSSEAVYDASALVAGSNLEG

AIQOILDMGGGTWDRDVIYIRVRAFNPERAVEXLYSGIPEQAEAPVPAPSPGQA

ANPLDOPPAAPAPASAGPNANPLDLEPQGLPDMGSNAAGNIDFLRTNQFOALR

AMVQSNPQIILQPMLOELGKNPHLMRLTIOHQADFLQLINEMEGEGMLIGHGQALIS

VTPEPRDAIERLEAMGPFRELIVLEVFACNKNNEELAAVYLLDHMEFE"

BASE COUNT 403 a 361 c 331 g 401 t

ORIGIN

alignment_scores:

Quality: 15.00

Ratio: 1.000

Percent Similarity: 100.000

Percent Identity: 100.000

alignment_block:

US-09-805-550-4 x DCRAD231 ..

Align seg 1/1 to: DCRAD231 from: 1 to: 1496

174 AsnAsnProGluArGAlaValGluTyLeuTySerGlyIlePro 188
 621 AACAAATCCAGAGAGAGCTGTGAGATCTTATTCCGGAATCCCT 665

seq_name: gb_pl:F309

seq_documentation_block:

LOCUS F309

DEFINITION Arabidopsis thaliana chromosome 1 BAC F309 sequence, complete

ACCESSION AC006341

VERSION AC006341.2 GI:4887257

KEYWORDS HTG.

SOURCE thale cress.

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 114498)

Vysotskaia,V.S., Schwartz,J., Yu,G., Toriumi,M., Lenz,C., Liu,S., Lee,J., Li,J., Kremetskaia,I., Liu,A., Luros,J., Gonzalez,A.,

Altati,H., Araujo,R., Chao,O., Conn,L., Conway,A.B., Dunn,P.,

Hansen,N., Huizar,L., Kim,C., Palm,C., Rowley,D., Shin,P.,

Walker,M., Davis,R.W., Ecker,J.R., Federspiel,N.A. and Theologis,A.

The sequence of BAC F309 from Arabidopsis thaliana chromosome 1

Unpublished (1999)

2 (bases 1 to 114498)

Theologis,A.

Direct Submission

Submitted (11-JAN-1999) Plant Gene Expression Center, 800 Buchanan

Street, Albany, CA 94710, USA

3 (bases 1 to 114498)

Theologis,A.

Direct Submission

Submitted (25-MAY-1999) Plant Gene Expression Center, 800 Buchanan

Street, Albany, CA 94710, USA

4 (bases 1 to 114498)

Theologis,A.

Direct Submission

Submitted (02-JUN-1999) Plant Gene Expression Center, 800 Buchanan

St., Albany, CA 94710, USA

On May 25, 1999 this sequence version replaced gi:4139327.

The sequence of BAC F309 from Arabidopsis thaliana chromosome 1.

FEATURES

source

1..114498

/organism="Arabidopsis thaliana"

/cultivar="Columbia"

/db_xref="taxon:3702"

/chromosome="1"

/clone="F309"

246..901

/gene="F309.1"

join(246..275,353..462,608..692,785..901)

/gene="F309.1"

/note="Similar to gb|Y12014 RAD23 protein isoform II from

daucus carota. This gene is probably cut off. Est

gb|AA651284 comes from this gene."

/codon_start=1

/evidence=not_experimental

/protein_id="AAD34676.1"

/db_xref="GI:4966345"

/translation="MVNSNPQIILQPMLOELGKNPDLRLIQENQAEFLQLNEPYEG

SDGDVDFEQDQPEMPHSVYVTPPEQESTERLEAMGFDRAYIEAFLECDRNEELAA

YLEHSADEP"

complement(1361..2444)

/gene="F309.2"

complement(join(1361..1387,1506..1625,1899..2444))

/gene="F309.2"

/note="ESTs gb|T04357 and gb|AA595092 come from this

gene."

/codon_start=1

/evidence=not_experimental

/protein_id="AAD34673.1"

/db_xref="GI:4966342"

/translation="GKINSKAEVAKSRKNAAEQKROTRKEBOYMRAGPASKA

VKREBEAEKKAETAKLKEAKRLAEOEKELKLPKDKANVTYVPTAEEL

IRREEDVYALAKKAEDSKKOTRMAAGDEYKAMLYVTNTNDDSLIEAHYDEALR

ITVSDNLVDRHPEKRLKASFRAVEVELPLKSEKPGILTYQYDLTKMKKSPDN

PLNQAAAAAN"

```

gene
3662..5562
/gene="F309.3"
join(3662..3916,4107..4475,4552..4669,4755..4978,
5053..5562)
/gene="F309.3"
/note="is a member of Pf00481 Protein phosphatase 2C
family."
/codon_start=1
/evidence-not_experimental
/protein_id="A034674.1"
/db_xref="GI:4966343"
/translation="MGICHSKIDKTRKRGATSTATTVEROSSGRLRRPRDLYSG
EISEIOVVGRLVNGSSEIACLYTOOGKGTNODAMLVNENCSRDYQCEQPOH
GPRCHVSKRYRDMLEPTLTQKTTSGTROSSKNGINAPTCVDEQMCCEQCEK
DEKLPPEMTPLKRALTKTCQOMDKELKMPRTINCPCSGTSTVVIKQGDVVGNG
DSRAVLATRDODNALVAVOLTIDLPKDPESARINHCGRVPALEDQEPVAVMLFN
SDSPGLAMARAFEDPCLKIDYGLSPDINHLERDOYIILATGDVADYLSKNEAD
IVASAPSDTAAVAVDQTAARMLKRYPTNDCQAVCLFEDTSGVGEVSEYN
HSHESESTYTTSSKRDADKEASPTNETPVMELEKEKEPESCHIESKTTLAC
ISVKDEBMSALGELTRVNSLSIPRFSELSNSSWRKWL"
complement(10192..11349)
/gene="F309.4"
complement(join(10192..10323,10518..10562,10642..10742,
10834..11039,11135..11349))
/gene="F309.4"
/note="ESTs gb|F15498, gb|H37515, gb|T41906, gb|T22448,
gb|W43356 and gb|T20739 come from this gene."
/evidence-not_experimental
/codon_start=1
/protein_id="A034675.1"
/db_xref="GI:4966344"
/translation="MASSSDSMKRAYNEALKSEINGATSESSSVYTPDQORAS
AIRKLTIRGNKLDLSQSLAEITHGKFRISKEKMRKRDVGNLRSKANQANALNNSN
FARSDILGPDIPDMSKRVGTMDNGIVQROVAREQDEGLEDEGTVMSTKHLA
LVASEEDLQTRLIDLDYHVDYDSRLRRVQSKLAVNMNMRSGSCSMGLLSVLGI
VGLAVYIWMLVKYM"
11987..13706
/gene="F309.5"
join(11987..12044,12084..12224,12408..12954,13133..13706)
/gene="F309.5"
/note="Contains two Pf101344 Kelch motif domains."
/evidence-not_experimental
/codon_start=1
/protein_id="A034677.1"
/db_xref="GI:4966346"
/translation="MPVSSVSLPQCNPEFLSDFSVSIIVNGVVGKVINFEAREWC
SKASRRGSEDTPTISRVRGSLIPDLALRCLAKLSHGHYHLECVSRGMHDLVRA
DYSCYKARNMGSMFLVLETERSKNOMVADPEADRMHPLPRTAVODGMHSGFACV
CVSNCLLVIGGCYAPSVSSPEPHQPVYTKDMREDDPEKKOMKVAISMRTPRFHACS
VSGKVVAGGRNLTSHRGIDSAEYDVPVADRMELPAMRPOKDCSLSTRGCFHYS
DOVGFABQNSSEVFNPRDNTWSTVEDVWPPSRAMQFQAVQMKNDRTYTIYDWGSLIK
TRDTDEGEVNVGSVPVLPNHPRELEAGYGFALRNELVYIGCVLWMBESSAGR
FDIVLRPVAVNCPLDRPLMWRPKPCIPAGSIIIGCVLSLESSSP"
complement(14717..17027)
/gene="F309.6"
complement(join(14717..15924,15993..16097,16178..17027))
/gene="F309.6"
/note="Similar to gb|A012423 wall-associated kinase 2
from Arabidopsis thaliana."
/codon_start=1
/evidence-not_experimental
/protein_id="A034678.1"
/db_xref="GI:4966347"
/translation="MGVDVARKPLVLMILLRICETAASTPLALRNCSDHGVNVPY
PFGIGGCKNKNWFEIYCKSSDQDPLILPRLRAVTSFNLDPSISVYNKFYIOS
PLKHSGCPNDGYSSSLNKGSPFTSENNKTPAVGCNNKAKMANTYQGLIVCEETC
GNEIRSYGKANTSCVGGKCCOMTIPPLDQVQDAVTEKREKOGQVADLFOFETS
GSLFTPELMEVSEYTTIELMRDLDSYMSKRVCLKFFSDSCCHNCEBGP
YIFGCGODIDECRDPHLNKGRCYVNVUSYCEKTPWPIILSGTSLSSGLLILFQMW
LICKANRRKVAOKRRFQVRNGGLLQDOTSFLHGSVNTKVFSSNDLNAVDRFNA
SRILGOGGQTVYKGLIDGMIVAVKSKALKSENEEFINEIILLSQIHNRNVKLI
GCLLETFPLVEYFIPNRILFDLHNPSEDFPMSWEKCIACEVADYLSHSAVS
IPYHRDVSTNLTDEKHKRAKVSDFGISVAIDTHTLTYYQGTIGYVDPETLQSN

```

```

gene
3662..5562
/gene="F309.3"
join(3662..3916,4107..4475,4552..4669,4755..4978,
5053..5562)
/gene="F309.3"
/note="is a member of Pf00481 Protein phosphatase 2C
family."
/codon_start=1
/evidence-not_experimental
/protein_id="A034674.1"
/db_xref="GI:4966343"
/translation="MGICHSKIDKTRKRGATSTATTVEROSSGRLRRPRDLYSG
EISEIOVVGRLVNGSSEIACLYTOOGKGTNODAMLVNENCSRDYQCEQPOH
GPRCHVSKRYRDMLEPTLTQKTTSGTROSSKNGINAPTCVDEQMCCEQCEK
DEKLPPEMTPLKRALTKTCQOMDKELKMPRTINCPCSGTSTVVIKQGDVVGNG
DSRAVLATRDODNALVAVOLTIDLPKDPESARINHCGRVPALEDQEPVAVMLFN
SDSPGLAMARAFEDPCLKIDYGLSPDINHLERDOYIILATGDVADYLSKNEAD
IVASAPSDTAAVAVDQTAARMLKRYPTNDCQAVCLFEDTSGVGEVSEYN
HSHESESTYTTSSKRDADKEASPTNETPVMELEKEKEPESCHIESKTTLAC
ISVKDEBMSALGELTRVNSLSIPRFSELSNSSWRKWL"
complement(10192..11349)
/gene="F309.4"
complement(join(10192..10323,10518..10562,10642..10742,
10834..11039,11135..11349))
/gene="F309.4"
/note="ESTs gb|F15498, gb|H37515, gb|T41906, gb|T22448,
gb|W43356 and gb|T20739 come from this gene."
/evidence-not_experimental
/codon_start=1
/protein_id="A034675.1"
/db_xref="GI:4966344"
/translation="MASSSDSMKRAYNEALKSEINGATSESSSVYTPDQORAS
AIRKLTIRGNKLDLSQSLAEITHGKFRISKEKMRKRDVGNLRSKANQANALNNSN
FARSDILGPDIPDMSKRVGTMDNGIVQROVAREQDEGLEDEGTVMSTKHLA
LVASEEDLQTRLIDLDYHVDYDSRLRRVQSKLAVNMNMRSGSCSMGLLSVLGI
VGLAVYIWMLVKYM"
11987..13706
/gene="F309.5"
join(11987..12044,12084..12224,12408..12954,13133..13706)
/gene="F309.5"
/note="Contains two Pf101344 Kelch motif domains."
/evidence-not_experimental
/codon_start=1
/protein_id="A034677.1"
/db_xref="GI:4966346"
/translation="MPVSSVSLPQCNPEFLSDFSVSIIVNGVVGKVINFEAREWC
SKASRRGSEDTPTISRVRGSLIPDLALRCLAKLSHGHYHLECVSRGMHDLVRA
DYSCYKARNMGSMFLVLETERSKNOMVADPEADRMHPLPRTAVODGMHSGFACV
CVSNCLLVIGGCYAPSVSSPEPHQPVYTKDMREDDPEKKOMKVAISMRTPRFHACS
VSGKVVAGGRNLTSHRGIDSAEYDVPVADRMELPAMRPOKDCSLSTRGCFHYS
DOVGFABQNSSEVFNPRDNTWSTVEDVWPPSRAMQFQAVQMKNDRTYTIYDWGSLIK
TRDTDEGEVNVGSVPVLPNHPRELEAGYGFALRNELVYIGCVLWMBESSAGR
FDIVLRPVAVNCPLDRPLMWRPKPCIPAGSIIIGCVLSLESSSP"
complement(14717..17027)
/gene="F309.6"
complement(join(14717..15924,15993..16097,16178..17027))
/gene="F309.6"
/note="Similar to gb|A012423 wall-associated kinase 2
from Arabidopsis thaliana."
/codon_start=1
/evidence-not_experimental
/protein_id="A034678.1"
/db_xref="GI:4966347"
/translation="MGVDVARKPLVLMILLRICETAASTPLALRNCSDHGVNVPY
PFGIGGCKNKNWFEIYCKSSDQDPLILPRLRAVTSFNLDPSISVYNKFYIOS
PLKHSGCPNDGYSSSLNKGSPFTSENNKTPAVGCNNKAKMANTYQGLIVCEETC
GNEIRSYGKANTSCVGGKCCOMTIPPLDQVQDAVTEKREKOGQVADLFOFETS
GSLFTPELMEVSEYTTIELMRDLDSYMSKRVCLKFFSDSCCHNCEBGP
YIFGCGODIDECRDPHLNKGRCYVNVUSYCEKTPWPIILSGTSLSSGLLILFQMW
LICKANRRKVAOKRRFQVRNGGLLQDOTSFLHGSVNTKVFSSNDLNAVDRFNA
SRILGOGGQTVYKGLIDGMIVAVKSKALKSENEEFINEIILLSQIHNRNVKLI
GCLLETFPLVEYFIPNRILFDLHNPSEDFPMSWEKCIACEVADYLSHSAVS
IPYHRDVSTNLTDEKHKRAKVSDFGISVAIDTHTLTYYQGTIGYVDPETLQSN

```

```

gene
3662..5562
/gene="F309.3"
join(3662..3916,4107..4475,4552..4669,4755..4978,
5053..5562)
/gene="F309.3"
/note="is a member of Pf00481 Protein phosphatase 2C
family."
/codon_start=1
/evidence-not_experimental
/protein_id="A034674.1"
/db_xref="GI:4966343"
/translation="MGICHSKIDKTRKRGATSTATTVEROSSGRLRRPRDLYSG
EISEIOVVGRLVNGSSEIACLYTOOGKGTNODAMLVNENCSRDYQCEQPOH
GPRCHVSKRYRDMLEPTLTQKTTSGTROSSKNGINAPTCVDEQMCCEQCEK
DEKLPPEMTPLKRALTKTCQOMDKELKMPRTINCPCSGTSTVVIKQGDVVGNG
DSRAVLATRDODNALVAVOLTIDLPKDPESARINHCGRVPALEDQEPVAVMLFN
SDSPGLAMARAFEDPCLKIDYGLSPDINHLERDOYIILATGDVADYLSKNEAD
IVASAPSDTAAVAVDQTAARMLKRYPTNDCQAVCLFEDTSGVGEVSEYN
HSHESESTYTTSSKRDADKEASPTNETPVMELEKEKEPESCHIESKTTLAC
ISVKDEBMSALGELTRVNSLSIPRFSELSNSSWRKWL"
complement(10192..11349)
/gene="F309.4"
complement(join(10192..10323,10518..10562,10642..10742,
10834..11039,11135..11349))
/gene="F309.4"
/note="ESTs gb|F15498, gb|H37515, gb|T41906, gb|T22448,
gb|W43356 and gb|T20739 come from this gene."
/evidence-not_experimental
/codon_start=1
/protein_id="A034675.1"
/db_xref="GI:4966344"
/translation="MASSSDSMKRAYNEALKSEINGATSESSSVYTPDQORAS
AIRKLTIRGNKLDLSQSLAEITHGKFRISKEKMRKRDVGNLRSKANQANALNNSN
FARSDILGPDIPDMSKRVGTMDNGIVQROVAREQDEGLEDEGTVMSTKHLA
LVASEEDLQTRLIDLDYHVDYDSRLRRVQSKLAVNMNMRSGSCSMGLLSVLGI
VGLAVYIWMLVKYM"
11987..13706
/gene="F309.5"
join(11987..12044,12084..12224,12408..12954,13133..13706)
/gene="F309.5"
/note="Contains two Pf101344 Kelch motif domains."
/evidence-not_experimental
/codon_start=1
/protein_id="A034677.1"
/db_xref="GI:4966346"
/translation="MPVSSVSLPQCNPEFLSDFSVSIIVNGVVGKVINFEAREWC
SKASRRGSEDTPTISRVRGSLIPDLALRCLAKLSHGHYHLECVSRGMHDLVRA
DYSCYKARNMGSMFLVLETERSKNOMVADPEADRMHPLPRTAVODGMHSGFACV
CVSNCLLVIGGCYAPSVSSPEPHQPVYTKDMREDDPEKKOMKVAISMRTPRFHACS
VSGKVVAGGRNLTSHRGIDSAEYDVPVADRMELPAMRPOKDCSLSTRGCFHYS
DOVGFABQNSSEVFNPRDNTWSTVEDVWPPSRAMQFQAVQMKNDRTYTIYDWGSLIK
TRDTDEGEVNVGSVPVLPNHPRELEAGYGFALRNELVYIGCVLWMBESSAGR
FDIVLRPVAVNCPLDRPLMWRPKPCIPAGSIIIGCVLSLESSSP"
complement(14717..17027)
/gene="F309.6"
complement(join(14717..15924,15993..16097,16178..17027))
/gene="F309.6"
/note="Similar to gb|A012423 wall-associated kinase 2
from Arabidopsis thaliana."
/codon_start=1
/evidence-not_experimental
/protein_id="A034678.1"
/db_xref="GI:4966347"
/translation="MGVDVARKPLVLMILLRICETAASTPLALRNCSDHGVNVPY
PFGIGGCKNKNWFEIYCKSSDQDPLILPRLRAVTSFNLDPSISVYNKFYIOS
PLKHSGCPNDGYSSSLNKGSPFTSENNKTPAVGCNNKAKMANTYQGLIVCEETC
GNEIRSYGKANTSCVGGKCCOMTIPPLDQVQDAVTEKREKOGQVADLFOFETS
GSLFTPELMEVSEYTTIELMRDLDSYMSKRVCLKFFSDSCCHNCEBGP
YIFGCGODIDECRDPHLNKGRCYVNVUSYCEKTPWPIILSGTSLSSGLLILFQMW
LICKANRRKVAOKRRFQVRNGGLLQDOTSFLHGSVNTKVFSSNDLNAVDRFNA
SRILGOGGQTVYKGLIDGMIVAVKSKALKSENEEFINEIILLSQIHNRNVKLI
GCLLETFPLVEYFIPNRILFDLHNPSEDFPMSWEKCIACEVADYLSHSAVS
IPYHRDVSTNLTDEKHKRAKVSDFGISVAIDTHTLTYYQGTIGYVDPETLQSN

```

```

gene
3662..5562
/gene="F309.3"
join(3662..3916,4107..4475,4552..4669,4755..4978,
5053..5562)
/gene="F309.3"
/note="is a member of Pf00481 Protein phosphatase 2C
family."
/codon_start=1
/evidence-not_experimental
/protein_id="A034674.1"
/db_xref="GI:4966343"
/translation="MGICHSKIDKTRKRGATSTATTVEROSSGRLRRPRDLYSG
EISEIOVVGRLVNGSSEIACLYTOOGKGTNODAMLVNENCSRDYQCEQPOH
GPRCHVSKRYRDMLEPTLTQKTTSGTROSSKNGINAPTCVDEQMCCEQCEK
DEKLPPEMTPLKRALTKTCQOMDKELKMPRTINCPCSGTSTVVIKQGDVVGNG
DSRAVLATRDODNALVAVOLTIDLPKDPESARINHCGRVPALEDQEPVAVMLFN
SDSPGLAMARAFEDPCLKIDYGLSPDINHLERDOYIILATGDVADYLSKNEAD
IVASAPSDTAAVAVDQTAARMLKRYPTNDCQAVCLFEDTSGVGEVSEYN
HSHESESTYTTSSKRDADKEASPTNETPVMELEKEKEPESCHIESKTTLAC
ISVKDEBMSALGELTRVNSLSIPRFSELSNSSWRKWL"
complement(10192..11349)
/gene="F309.4"
complement(join(10192..10323,10518..10562,10642..10742,
10834..11039,11135..11349))
/gene="F309.4"
/note="ESTs gb|F15498, gb|H37515, gb|T41906, gb|T22448,
gb|W43356 and gb|T20739 come from this gene."
/evidence-not_experimental
/codon_start=1
/protein_id="A034675.1"
/db_xref="GI:4966344"
/translation="MASSSDSMKRAYNEALKSEINGATSESSSVYTPDQORAS
AIRKLTIRGNKLDLSQSLAEITHGKFRISKEKMRKRDVGNLRSKANQANALNNSN
FARSDILGPDIPDMSKRVGTMDNGIVQROVAREQDEGLEDEGTVMSTKHLA
LVASEEDLQTRLIDLDYHVDYDSRLRRVQSKLAVNMNMRSGSCSMGLLSVLGI
VGLAVYIWMLVKYM"
11987..13706
/gene="F309.5"
join(11987..12044,12084..12224,12408..12954,13133..13706)
/gene="F309.5"
/note="Contains two Pf101344 Kelch motif domains."
/evidence-not_experimental
/codon_start=1
/protein_id="A034677.1"
/db_xref="GI:4966346"
/translation="MPVSSVSLPQCNPEFLSDFSVSIIVNGVVGKVINFEAREWC
SKASRRGSEDTPTISRVRGSLIPDLALRCLAKLSHGHYHLECVSRGMHDLVRA
DYSCYKARNMGSMFLVLETERSKNOMVADPEADRMHPLPRTAVODGMHSGFACV
CVSNCLLVIGGCYAPSVSSPEPHQPVYTKDMREDDPEKKOMKVAISMRTPRFHACS
VSGKVVAGGRNLTSHRGIDSAEYDVPVADRMELPAMRPOKDCSLSTRGCFHYS
DOVGFABQNSSEVFNPRDNTWSTVEDVWPPSRAMQFQAVQMKNDRTYTIYDWGSLIK
TRDTDEGEVNVGSVPVLPNHPRELEAGYGFALRNELVYIGCVLWMBESSAGR
FDIVLRPVAVNCPLDRPLMWRPKPCIPAGSIIIGCVLSLESSSP"
complement(14717..17027)
/gene="F309.6"
complement(join(14717..15924,15993..16097,16178..17027))
/gene="F309.6"
/note="Similar to gb|A012423 wall-associated kinase 2
from Arabidopsis thaliana."
/codon_start=1
/evidence-not_experimental
/protein_id="A034678.1"
/db_xref="GI:4966347"
/translation="MGVDVARKPLVLMILLRICETAASTPLALRNCSDHGVNVPY
PFGIGGCKNKNWFEIYCKSSDQDPLILPRLRAVTSFNLDPSISVYNKFYIOS
PLKHSGCPNDGYSSSLNKGSPFTSENNKTPAVGCNNKAKMANTYQGLIVCEETC
GNEIRSYGKANTSCVGGKCCOMTIPPLDQVQDAVTEKREKOGQVADLFOFETS
GSLFTPELMEVSEYTTIELMRDLDSYMSKRVCLKFFSDSCCHNCEBGP
YIFGCGODIDECRDPHLNKGRCYVNVUSYCEKTPWPIILSGTSLSSGLLILFQMW
LICKANRRKVAOKRRFQVRNGGLLQDOTSFLHGSVNTKVFSSNDLNAVDRFNA
SRILGOGGQTVYKGLIDGMIVAVKSKALKSENEEFINEIILLSQIHNRNVKLI
GCLLETFPLVEYFIPNRILFDLHNPSEDFPMSWEKCIACEVADYLSHSAVS
IPYHRDVSTNLTDEKHKRAKVSDFGISVAIDTHTLTYYQGTIGYVDPETLQSN

```

```

gene
3662..5562
/gene="F309.3"
join(3662..3916,4107..4475,4552..4669,4755..4978,
5053..5562)
/gene="F309.3"
/note="is a member of Pf00481 Protein phosphatase 2C
family."
/codon_start=1
/evidence-not_experimental
/protein_id="A034674.1"
/db_xref="GI:4966343"
/translation="MGICHSKIDKTRKRGATSTATTVEROSSGRLRRPRDLYSG
EISEIOVVGRLVNGSSEIACLYTOOGKGTNODAMLVNENCSRDYQCEQPOH
GPRCHVSKRYRDMLEPTLTQKTTSGTROSSKNGINAPTCVDEQMCCEQCEK
DEKLPPEMTPLKRALTKTCQOMDKELKMPRTINCPCSGTSTVVIKQGDVVGNG
DSRAVLATRDODNALVAVOLTIDLPKDPESARINHCGRVPALEDQEPVAVMLFN
SDSPGLAMARAFEDPCLKIDYGLSPDINHLERDOYIILATGDVADYLSKNEAD
IVASAPSDTAAVAVDQTAARMLKRYPTNDCQAVCLFEDTSGVGEVSEYN
HSHESESTYTTSSKRDADKEASPTNETPVMELEKEKEPESCHIESKTTLAC
ISVKDEBMSALGELTRVNSLSIPRFSELSNSSWRKWL"
complement(10192..11349)
/gene="F309.4"
complement(join(10192..10323,10518..10562,10642..10742,
10834..11039,11135..11349))
/gene="F309.4"
/note="ESTs gb|F15498, gb|H37515, gb|T41906, gb|T22448,
gb|W43356 and gb|T20739 come from this gene."
/evidence-not_experimental
/codon_start=1
/protein_id="A034675.1"
/db_xref="GI:4966344"
/translation="MASSSDSMKRAYNEALKSEINGATSESSSVYTPDQORAS
AIRKLTIRGNKLDLSQSLAEITHGKFRISKEKMRKRDVGNLRSKANQANALNNSN
FARSDILGPDIPDMSKRVGTMDNGIVQROVAREQDEGLEDEGTVMSTKHLA
LVASEEDLQTRLIDLDYHVDYDSRLRRVQSKLAVNMNMRSGSCSMGLLSVLGI
VGLAVYIWMLVKYM"
11987..13706
/gene="F309.5"
join(11987..12044,12084..12224,12408..12954,13133..13706)
/gene="F309.5"
/note="Contains two Pf101344 Kelch motif domains."
/evidence-not_experimental
/codon_start=1
/protein_id="A034677.1"
/db_xref="GI:4966346"
/translation="MPVSSVSLPQCNPEFLSDFSVSIIVNGVVGKVINFEAREWC
SKASRRGSEDTPTISRVRGSLIPDLALRCLAKLSHGHYHLECVSRGMHDLVRA
DYSCYKARNMGSMFLVLETERSKNOMVADPEADRMHPLPRTAVODGMHSGFACV
CVSNCLLVIGGCYAPSVSSPEPHQPVYTKDMREDDPEKKOMKVAISMRTPRFHACS
VSGKVVAGGRNLTSHRGIDSAEYDVPVADRMELPAMRPOKDCSLSTRGCFHYS
DOVGFABQNSSEVFNPRDNTWSTVEDVWPPSRAMQFQAVQMKNDRTYTIYDWGSLIK
TRDTDEGEVNVGSVPVLPNHPRELEAGYGFALRNELVYIGCVLWMBESSAGR
FDIVLRPVAVNCPLDRPLMWRPKPCIPAGSIIIGCVLSLESSSP"
complement(14717..17027)
/gene="F309.6"
complement(join(14717..15924,15993..16097,16178..17027))
/gene="F309.6"
/note="Similar to gb|A012423 wall-associated kinase 2
from Arabidopsis thaliana."
/codon_start=1
/evidence-not_experimental
/protein_id="A034678.1"
/db_xref="GI:4966347"
/translation="MGVDVARKPLVLMILLRICETAASTPLALRNCSDHGVNVPY
PFGIGGCKNKNWFEIYCKSSDQDPLILPRLRAVTSFNLDPSISVYNKFYIOS
PLKHSGCPNDGYSSSLNKGSPFTSENNKTPAVGCNNKAKMANTYQGLIVCEETC
GNEIRSYGKANTSCVGGKCCOMTIPPLDQVQDAVTEKREKOGQVADLFOFETS
GSLFTPELMEVSEYTTIELMRDLDSYMSKRVCLKFFSDSCCHNCEBGP
YIFGCGODIDECRDPHLNKGRCYVNVUSYCEKTPWPIILSGTSLSSGLLILFQMW
LICKANRRKVAOKRRFQVRNGGLLQDOTSFLHGSVNTKVFSSNDLNAVDRFNA
SRILGOGGQTVYKGLIDGMIVAVKSKALKSENEEFINEIILLSQIHNRNVKLI
GCLLETFPLVEYFIPNRILFDLHNPSEDFPMSWEKCIACEVADYLSHSAVS
IPYHRDVSTNLTDEKHKRAKVSDFGISVAIDTHTLTYYQGTIGYVDPETLQSN

```

KEYWORDS	HTG.
SOURCE	thale cress.
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryotes: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; Rosidae: eucosids II; Brassicales; Brassicaceae; Arabidopsis.
AUTHORS	1 (bases 1 to 80442)
REFERENCE	Liu,S.X., Yu,G., Sakano,H., Thaler,A., Lee,J.M., Lenz,C., Pham,P., Toriumi,M., Chih,C., Chio,J., Choi,E., Chung,M., Gonzalez,A.,
JOURNAL	Howm,B., Koo,T., Li,J., Liu,A., Vaysberg,M., Alcafi,H., Brooks,S., Buehler,E., Chao,O., Conn,L., Conway,A., Hansen,N., Johnson-Hopson,C., Khan,S., Kim,C., Lam,B., Nguyen,M., Palm,C., Shih,P., Tambunga,G., Davis,R.W., Ecker,J.R., Federspiel,N.A. and Theologis,A.
TITLE	The sequence of BAC T24D18 from Arabidopsis thaliana chromosome 1
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 80442)
AUTHORS	Theologis,A.
JOURNAL	Direct Submission
REFERENCE	Submitted (28-SEP-1999) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
TITLE	3 (bases 1 to 80442)
JOURNAL	Theologis,A.
REFERENCE	Direct Submission
AUTHORS	Submitted (06-NOV-1999) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
JOURNAL	4 (bases 1 to 80442)
REFERENCE	Theologis,A.
AUTHORS	Direct Submission
JOURNAL	Submitted (16-DEC-1999) Plant Gene Expression Center, 800 Buchanan St., Albany, CA 94710, USA
COMMENT	On Nov 6, 1999 this sequence version replaced gi:5931388. The sequence of BAC T24D18 from Arabidopsis thaliana chromosome 1. The sequence does not represent the sequence of the entire insert of the unique sequence of the clone. It is shorter by 1381 bp because we submit only the unique sequence of the clone. However, in order to facilitate the joining of overlapping clones in the future for creation of larger contigs, we provide small overlaps (200 bp) between overlapping submitted clones. The '3' end of this sequence overlaps by 200 bp the '5' end of the sequence of the clone F309.
FEATURES	Location/Qualifiers
SOURCE	1..80442
gene	/organism="Arabidopsis thaliana"
CDS	/cultivar="Columbia"
gene	/db_xref="taxon:3702"
CDS	/chromosome="1"
gene	/clone="T24D18"
CDS	1202..3974
gene	/gene="T24D18.1"
CDS	John(1202..2102,2184..2365,2479..2568,2682..2969,3324..3536,3744..3974)
gene	/gene="T24D18.1"
CDS	/note="Contains similarity to gb AF136530 transcripional regulator from Zea mays. ESTs gb F14071, gb Z26823, gb A1998935 come from this gene."
gene	/codon_start=1
CDS	/evidence=not experimental
gene	/protein_id="AA18488.1"
CDS	/db_xref="GI:6587797"
gene	/translation="MSISDEAEISEIEDYSETPYRLRDGTYYKVGOLCPCPCAGKKADKYKKEIYAHATGVSGATSAALOKAHNLAMLELNLGVAEPVPRPVVPOLDTEPNPHNYVPMWGIYVPLKEADDELLDSYMLQTSKFEPIVMAHWPDDLTGVYIAKNGMGSPGATATELEKEGSGSKKEMTERSGSGSKAYWCARADPESGCPIGEYISKRGOLQATYSDISKVQDNRYLLEISDMITANTNEDLNVOYSINRFLMSYLVDEKKNLHOFADERTKKMOMSRHLOKLYDEKLSNELDRMHODI ESRKQLEKHEALTELDROKIDEDRRSDMMKSLQASLRQKADSVSLVDEHORRDKKALNNKILLLEKOLDTKOTLEMEIOLGKLGQVMNHAGDDDEAVQKKKEANDELDKKAEEGLEGSMSVLTKEKROGNDIOLARKKLINGLGLGAETFDIGYKRGELDEKPLDQCKLRISANEAVAEATITCSWENLRKPSMOPKHEGTGGAEEVYEDDEDLKKREMGKEVHNNAVKTALVENETNASGRITTPELMFKESGRKATLKEVITTFISNSDIKLKKRRK"
gene	/gene="T24D18.2"
gene	5213..6073

CDS 5213..6073
/gene="T24D18.2"
/note="Similar to g11G0809 CCR4-associated factor 1 (CAF1) from Mus musculus. EST gb|226822 comes from this gene."
/codon_start=1
/evidence=not_experimental
/protein_id="AAI18489.1"
/db_xref="GI:6587798"
/translation="MSQAPNPDEEDPTIEIREVWNHNLDEOMALIEOSIDDFPVAND
TEPPGIVCKVTANPNPNPYSIHEYDYTLKANVNLKIQGLTISDERGNLPTGC
TNKOCIMQFNREFRANYSIDFALDSIELYKSAIDLEKNECGDARFPELMSGCV
VLNDKIHMTWHEHCGYDGYLLKLLSGEILPELISDPEDMEKEFPVYDLYKLMGRT
NLXGLEKIAELLGKVRKVGISHQAGSDSLTLTKFTKKEFFTGSLIKTSGFLGID
NFKLLTGSKN"
/gene="T24D18.3"
join(6953..6960,7349..7649,7771..7836,7933..7992)
/gene="T24D18.3"
/note="Strong similarity to gblAF067732 ribosomal protein
S12 from Hordeum vulgare. ESTs gb|1741772, gb|1742570,
gb|A1999345, gb|1720784, gb|F20068 come from this gene."
/codon_start=1
/evidence=not_experimental
/protein_id="AAI18490.1"
/db_xref="GI:6587799"
/translation="MSGDENAPVVPVPPAEPAPAIPEDDMLTALLETTLKARAYGV
VRGLHECAKLIKERVAKQVVLADCNQPDVVKLAKACDHEVRLLVPSAKTLGEWA
GLCKIDSEGNARVAGVCSCLVVKDFEETALISVKNHIASQ"
8923..13301
/gene="T24D18.4"
join(8923..9036,9346..9549,9665..9846,9939..10065,
10150..10347,10469..10519,10604..11311,11396..11493,
11579..11652,11736..11787,11877..11963,12052..12095,
12184..12258,12341..13301)
/gene="T24D18.4"
/codon_start=1
/evidence=not_experimental
/product="T24D18.4"
/protein_id="AAI18491.1"
/db_xref="GI:6587800"
/translation="MGPLVEATFEKSEQALTDAAENLKPFSSTDAITLLEVESIL
ATVEODLSSVOKALHPMRAIVSADLRNPDSDVRSVSCLEIMRTLAAPAPVND
EOMKDIPOVTEAFKLIADASSRYKAEVILTVAKVRSILWLDLDCDVLNEMO
RFLKTRPDHPOLVLYMETIMITVIDESEVPMDELITLVTKKDSOVSAATL
VEKVLSCCTCKLQPCIMEALKSSGSLDMVSPVSSICQSEFAITQAHNDKRDNA
DEKISGOVVPVSDLEKLNGLSRKGTSRSGARAGTTRANDDEVITANESEST
DAETASGSTRKRGMRKPSLAMPREGYSFTKSSSKVOEKELGDSLSGVAKKVPLPS
KVGQTNQSVVISLSSSGARATRSKRSTRMEETDDVSVATOPAKKOTVKTNPAAK
EDLTNRNVKHEDEGKITCKSSKKEKADNGLAKTSAKKPLAETMMVKRSGKLVHSDAK
KNSGASMDPTLPQSSSKSKKDSRATTPATKSEQAPKSHPMKRIAGEVESNTNE
LGEELVGRNVWVPLDKRFEYGVIKSYCVKTMQVTVSDGVEELNKEFEKITE
DKSSASEDEDLLESTPLSAFIQREKSKRKIVSNVESSSPEVSSQOKMKKDS
VDSIKOTRKTKALKAVSNPESTGTNKLKSLKNGEEDTGRGRKMOXQYAOAH
RKTEKDCDEOEDLTKDEEDSLKIGKESAPRMDHOLPPNNHVEPTKCOXDEOA
AKEPRAESTNGCEPNAEPETDCKEHSKLEKPAEKRSQSEEDBAKKEPAELKTQSE
NOBAKELTAERKTDEEHRKVADEVEQSKDEKYNVEPEAGEQKSEENAPKPKYK
EKESEAKEQYADTKLIEKEDMSKTEEDIKETYSIPETGKGVNEAEDDDOVIKEL
EESSDRAEVSTVLEVP"
14106..17166
/gene="T24D18.5"
join(14106..14238,14360..14514,15043..15228,15405..15757,
16959..17166)
/gene="T24D18.5"
/note="Strong similarity to cinamoyl CoA reductase
gbl12960364 from Populus balsamifera. ESTs gblN95902,
gblA1992693, gblA1995837 come from this gene."
/codon_start=1
/evidence=not_experimental
/protein_id="AAI18492.1"
/db_xref="GI:6587801"
/translation="MPVDVAPGAKTVCVTGAGGYIASNIVKILLEGYTVKGTVRNP
DPRKNHMLRELBGKERRILICKADLDQYALKAKAIDGCGVPHFAPSVYDDPRQMRP

gene
CDS
complement(join(17453..17770,17863..17958,18041..18172,
18248..18352,18449..18622,18698..18767,18839..19020,
19098..19207,19342..19431,19565..19643,19731..19771,
19858..19936,20225..20317))
/gene="T24D18.6"
/note="Strong similarity to g112130080 Nramp1 protein from
Oryza sativa and contains a Pf01566 Nramp domain."
/codon_start=1
/evidence=not_experimental
/protein_id="AAI18493.1"
/db_xref="GI:6587802"
/translation="MAETRSQSNRSINSPLIENSQNLIVPEKSKMKNFSTYLP
GFLVSLAYIDPGRFEDLQSGAQRYKELMILIVASCAALVQSILANLGVGKHLA
EHCRAEYSKVPNFMVWVAVIALVGFALNMFNIPVMTGVLTLGLSTILAL
QOYGIKLEFLIAFLVETLALCFEVELHYKPPKEVLYGLFVPOLGNGATGLAISL
LGAMVMPHNLFLSALVLSRKIPRSYTGKREACRYVLESGLAMVAFLVNSVYS
GAYCNSDLSPEPRAQODDLNKASFLRNVYKSSKRFALAILASGSSITTCY
AGQYMOGFLDLRLERPLRNFLRCLAITPSLIALGSGAGAKLITIASMLISEL
PFALVPLKFTSSKTKGSHANSVLVSSVWIIIGLIMGINITYLVSSFKLLHSHM
NLVAIVFLGVLPSTKTVGIALTVLAALVSLRKNRSSSTHFLDFNSQTEETLPREDIAN
MOLPNRAVIGDNL"
complement(21795..23745)
/gene="T24D18.7"
complement(join(21795..22106,22210..22351,22466..22549,
22637..22694,23283..23745))
/gene="T24D18.7"
/note="Contains similarity to g1112785
DNA-3-methyladenine glycosylase I from Escherichia coli."
/codon_start=1
/evidence=not_experimental
/protein_id="AAI18494.1"
/db_xref="GI:6587803"
/translation="MSVPRFRSVNSDEREFRSVLGPTGNKLRKPPGMLEKPMMEK
TIIDSKDEKAKPPTTPASPRTLTKOCSSLSILRKNSAMTASYSDDSSCESSPL
SVASSSSCKKVVRRSGSVSTRLSKQKEEKVSGCFFADGRCAWITPKADPCYVA
FHEBNGVPVPHDOKLPFLCLISGALAELSMTITILSRHILKREVPDFVAYAEIND
KLTLPAPTAISLISEVKIRSLIDNSHVKIILAECSLKAKYMNPNVNNKPTQSOFRY
QROVPVKTSAEFLISDLVRRGRSVSPVIVSFMQAAGLTNDHLIGCFRYODCCYDA
ETITTTKAKKKNRESDK"
24565..26450
/gene="T24D18.8"
join(24565..24989,25088..25238,25326..25516,25589..25846,
25946..26450)
alignment_scores:
Quality: 14.00 Length: 14
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-805-550-4 x T24D18 ..
Align seg 1/1 to: T24D18 from: 1 to: 80442
167 ARGAlaLeuAAGAlaAlaIaTyTAsnaSnProGuaIgaAlaVal 180
|||||
79856 CGTGCACCTTCGTGACATATTAACAACCTGAGAGACAGTG 79897
|||||
